Contribution ID: **205** Type: **Oral**

Optimising HEP parameter fits through MC weight derivative regression

Thursday, 7 November 2019 14:45 (15 minutes)

HEP event selection is traditionally considered a binary classification problem, involving the dichotomous categories of signal and background. In distribution fits for particle masses or couplings, however, signal events are not all equivalent, as the signal differential cross section has different sensitivities to the measured parameter in different regions of phase space. In this talk, I describe a mathematical framework for the evaluation and optimization of HEP parameter fits, where this sensitivity is defined on an event-by-event basis, and for MC events it is modeled in terms of their MC weight derivatives with respect to the measured parameter. Minimising the statistical error on a measurement implies the need to resolve (i.e. separate) events with different sensitivities, which ultimately represents a non-dichotomous classification problem. Since MC weight derivatives are not available for real data, the practical strategy I suggest consists in training a regressor of weight derivatives against MC events, and then using it as an optimal partitioning variable for 1-dimensional fits of data events. I discuss some features and limitations of this approach and I show examples from a simple toy model. This research is an extension of the work I presented at CHEP2018: in particular, event-by-event sensitivities allow the exact computation of the "FIP" ratio (a metric in [0,1]) between the Fisher information obtained from an analysis and the maximum information that could possibly be obtained with an ideal detector. Using this expression, I discuss the relationship between FIP and two metrics commonly used in meteorology (Brier score and MSE), and the importance of "sharpness" in both domains. I finally point out that HEP distribution fits should be optimized and evaluated using probabilistic metrics (like FIP or MSE), whereas ranking metrics (like AUC) or threshold metrics (like "accuracy") are of limited relevance for these specific problems.

Consider for promotion

No

Primary author: VALASSI, Andrea (CERN)

Presenter: VALASSI, Andrea (CERN)

Session Classification: Track 6 –Physics Analysis

Track Classification: Track 6 –Physics Analysis